

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 07:12:43 : Search time 2325.7 Seconds  
(without alignments)  
992.189 Million cell updates/sec

Title: US-09-303-518d-127  
Perfect score: 1344  
Sequence: 1 atgattaaatcaaaaagg.....conttgagaaggaagctga 1344

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

N\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	1327	98.7	1344	20	AAZ12027	Neisseria meningit
2	1229.4	91.5	1344	20	AAZ12026	Neisseria meningit
3	1229.4	91.5	44608	21	AAA81495	N. meningitidis pa
4	1229.4	91.5	349980	20	AAF21607	Neisseria meningit
5	1165.4	86.7	1344	20	AAZ12028	Neisseria gonorrhoe
6	527.2	39.2	1830121	17	NAF42063	Haemophilus influe
7	431.6	32.1	474	20	AAZ12025	Neisseria meningit
8	431.6	32.1	474	21	AAA81335	N. meningitidis Me
9	351	26.1	363	21	AAZ54035	Neisseria meningit

C	10	333.6	24.8	96109	22	AAF28548	Genomic fragment #
C	11	317.4	23.6	363	21	AAZ54034	Neisseria meningit
C	12	291.8	21.7	363	21	AAZ54033	Neisseria gonorrhoe
C	13	122.2	9.1	1353	20	AAZ91657	Porphyromonas ging
C	14	122.2	9.1	1362	20	AAZ91536	Porphyromonas ging
C	15	39.2	2.9	5059	20	AAZ84332	Stealth virus nucl
C	16	38.2	2.9	9762	21	AAZ48249	Rubella virus Cend
C	17	38.2	2.8	65140	22	AAZ17184	Rubella virus Cend
C	18	38.2	2.8	125401	22	AAZ17186	Streptomyces nous
C	19	37.2	2.8	9759	19	AAV34766	Rubella virus RA27
C	20	37	2.8	763	20	AAZ09101	Mistletoe lectin A
C	21	37	2.8	1598	20	AAZ09100	Mistletoe lectin D
C	22	36.6	2.7	1401	21	AAZ95378	Chlamydia pneumoni
C	23	36.4	2.7	1325	22	AAZ08837	Human G-protein co
C	24	36.4	2.7	4096	22	AAZ08837	Corynebacterium th
C	25	36	2.7	762	20	AAZ09104	Mistletoe lectin A
C	26	36	2.7	3230	22	AAH16235	Human cDNA sequenc
C	27	36	2.7	92934	21	AAA81473	N. meningitidis pa
C	28	36	2.7	172325	21	AAF21613	Neisseria meningit
C	29	36	2.7	837096	21	AAZ1489	N. meningitidis pa
C	30	35.8	2.7	699	22	AAZ97005	Mycobacterium tube
C	31	35.8	2.7	4403765	22	AAZ99683	Mycobacterium tube
C	32	35.8	2.7	4411529	22	AAZ99682	f-Therien rubella
C	33	35.6	2.6	2558	21	AAZ08045	Infectious rubella
C	34	35.6	2.6	9757	16	AAZ07686	Infectious rubella
C	35	35.6	2.6	9759	18	AAZ08642	Rubella virus CDNA
C	36	35.6	2.6	9759	21	AAZ08642	Pseudomonas aerugi
C	37	35.4	2.6	2727	23	AAZ54113	cDNA #195 encoding
C	38	35	2.6	296	23	AAZ57519	Nitrosomonas dnaK
C	39	35	2.6	1935	19	AAZ23811	DNA encoding novel
C	40	34.6	2.6	323	23	AAZ86095	Human breast and o
C	41	34.6	2.6	1139	21	AAF21735	DNA encoding novel
C	42	34.6	2.6	1689	23	AAZ86098	DNA encoding novel
C	43	34.6	2.6	6237	23	AAZ76357	DNA encoding novel
C	44	34.4	2.6	1050	23	AAZ73416	DNA encoding novel
C	45	34.4	2.6	1050	23	AAZ80144	Human polynucleoti
C	46	34.4	2.6	1774	22	AAZ52900	Endotoxin Cryg gen
C	47	34.4	2.6	2165	18	AAZ59702	N. meningitidis pa
C	48	34.4	2.6	18974	21	AAZ1485	Neisseria meningit
C	49	34.4	2.6	34980	21	AAF21607	N. meningitidis B
C	50	34.4	2.6	1437688	21	AAZ81490	Human protein HP03
C	51	34.2	2.5	1116	22	AAF28683	Human G protein-co
C	52	34.2	2.5	1116	24	ABA02176	DNA encoding a p2v
C	53	34.2	2.5	1119	21	AAZ64367	Human orphan G pro
C	54	34.2	2.5	1119	21	AAZ01119	Human G protein co
C	55	34.2	2.5	1119	22	AAZ49504	Human GTP-binding
C	56	34.2	2.5	1119	22	AAZ02585	Human G-protein co
C	57	34.2	2.5	1119	22	AAZ02585	Human G-protein co
C	58	34.2	2.5	1119	22	AAZ02585	Human DNA for pote
C	59	34.2	2.5	1119	24	AAZ98045	Human cDNA encodin
C	60	34.2	2.5	1237	24	AAZ98085	Human protein HP03
C	61	34.2	2.5	1560	24	AAZ19414	Human G-protein co
C	62	34.2	2.5	1720	22	AAZ28693	Human CON217 G pro
C	63	34.2	2.5	2444	22	AAZ26369	Human G protein co
C	64	34.2	2.5	2444	22	AAZ26369	Human P2Y1i DNA
C	65	34.2	2.5	2480	22	AAZ06509	Genomic sequence #
C	66	34.2	2.5	2559	21	AAZ95039	Human digestive sy
C	67	34.2	2.5	3180	22	AAZ25830	Protonibacterium
C	68	34.2	2.5	3789	22	AAZ66883	C glutamicum codin
C	69	34.2	2.5	15783	22	AAZ39803	C glutamicum codin
C	70	34.2	2.5	15783	22	AAZ39803	N. meningitidis pa
C	71	34.2	2.5	25365	23	AAZ95058	Platenolide syntha
C	72	34.2	2.5	349980	22	AAZ68529	N. meningitidis pa
C	73	34.2	2.5	349980	22	AAZ68529	Neisseria meningit
C	74	34	2.5	10577	21	AAZ81494	Protonibacterium
C	75	34	2.5	44377	18	AAZ78508	P. putida oxygenas
C	76	34	2.5	44377	18	AAZ78508	DNA encoding a su
C	77	34	2.5	69936	21	AAZ1479	
C	78	33.8	2.5	18796	23	AAZ21610	
C	79	33.8	2.5	29379	23	AAZ59517	
C	80	33.8	2.5	29379	23	AAZ59510	
C	81	33.6	2.5	621	22	AAZ26372	
C	82	33.6	2.5	10732	21	AAZ10594	

	83	33.4	2.5	4590	22	AAH24065	Yeast AOD9604-asso
	83	33.4	2.5	4590	22	AAH24065	Drosophila melanog
	84	33.2	2.5	1544	23	ABL06725	Drosophila melanog
	85	33.2	2.5	3866	23	ABL06724	Drosophila melanog
	86	33	2.5	245	21	AAA44957	Human secreted exp
	87	33	2.5	1311	21	AAAC4295	Arabidopsis thalia
	88	33	2.5	5416	19	AAV07381	Neisseria meningit
	89	32.8	2.4	876	23	ABL21113	Drosophila melanog
	90	32.8	2.4	1137	20	ABL16599	Drosophila melanog
	91	32.8	2.4	1596	20	AAZ09103	Mistletoe lectin I
	92	32.8	2.4	4399	23	ABL16598	Drosophila melanog
	93	32.8	2.4	4309	23	ABL21112	Drosophila melanog
	94	32.6	2.4	1451	21	AAF14884	Trichodema ressel
	95	32.6	2.4	14041	22	AAH48024	Internal control B
	96	32.4	2.4	754	22	AAH00214	Stenotrophomonas m
	97	32.4	2.4	1191	22	AAH66975	C glutamicum codin
	98	32.4	2.4	1400	13	AAQ22689	Sequence encoding
	99	32.4	2.4	2394	21	AAAL5156	DNA encoding outer
	100	32.4	2.4	2394	22	AA507379	Neisseria gonorrhoe

## ALIGNMENTS

RESULT	1	
AAAZ12027		
ID	AAZ12027 standard; DNA; 1344 BP.	
XX		
AC	AAZ12027;	
XX		
DT	08-OCT-1999 (first entry)	
XX		
DE	Neisseria meningitidis strain A complete ORF22 sequence.	
XX		
DE	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
XX	treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	W09924578-A2.	
XX		
PD	20-MAY-1999.	
XX		
PP	09-OCT-1998; 98WO-IB01665.	
XX		
PR	01-SEP-1998; 98GB-0019016.	
XX		
PR	06-NOV-1997; 97GB-0023516.	
XX		
PR	14-NOV-1997; 97GB-0024190.	
XX		
PR	18-NOV-1997; 97GB-0024386.	
XX		
PR	27-NOV-1997; 97GB-0025158.	
XX		
PR	10-DEC-1997; 97GB-0026147.	
XX		
PA	14-JAN-1998; 98GB-0000759.	
XX		
PA	(CHIR-) CHIRON SPA.	
XX		
PI	Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;	
XX		
DR	WPI; 1999-327407/27.	
XX		
PT	P-PSDB; AAY38562.	
XX		
PT	proteins from Neisseria meningitidis and N. gonorrhoeae useful for	
XX	diagnosis, treatment and prevention of infection	
XX		
PS	Claim 9; Page 123; 524pp; English.	
XX		
CC	Nucleotide sequences AAZ11972-212358 represent open reading frames	
XX	(ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode	
CC	antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their	
CC	fragments, their nucleic acids and antibodies are used for diagnosis,	
CC	prevention (as vaccines) or treatment of Neisseria infections,	
CC	such as meningitis, septicaemia and gonorrhea. Both organisms	
CC	are closely related. Fragments of the nucleic acids are useful	
CC	as hybridisation probes and antisense reagents.	
XX		

Db 961 cacaatcagattccgttaccgaagaagccgagcaagagctgttcggtggtg 1020  
QY 1021 ccgagccggacaaatctccatccacgctacgacccctggtccatttctgaaacaaa 1080  
Db 1021 ccgagccggacaaatctccatccacgctacgacccctggtccatttctgaaacaaa 1080  
QY 1081 ctcttcagttccacagccgctcaacggtggtgagccgctggtccgattggtact 1140  
Db 1081 ctcttcagttccacagccgctcaacggtggtgagccgctggtccgattggtact 1140  
QY 1141 tacgagccgtaataccgctgagacactcctgctacccctggttgcgagattatgctc 1200  
Db 1141 tacgagccgtaataccgctgagacactcctgctacccctggttgcgagattatgctc 1200  
QY 1201 ggcgtaccgagccgctgagacattggttgcgttggaattggagcaagaagacctgcct 1260  
Db 1201 ggcgtaccgagccgctgagacattggttgcgttggaattggagcaagaagacctgcct 1260  
QY 1261 ttgtcagcttcgtgctgcccgggcaaatcacgaatggtggttgcgttaaggtgctg 1320  
Db 1261 ttgtcagcttcgtgctgcccgggcaaatcacgaatggtggttgcgttaaggtgctg 1320  
QY 1321 gaaacnttgagaaggaaggtga 1344  
Db 1321 gaaacnttgagaaggaaggtga 1344

## RESULT 2

AAZ12026  
ID AAZ12026 standard; DNA; 1344 BP.  
XX  
AC AAZ12026;  
XX  
DT 08-OCT-1999 (first entry)  
XX  
DE Neisseria meningitidis complete ORF22 sequence.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.  
OS Neisseria meningitidis.  
XX  
PN WO924578-A2.  
XX  
PD 20-MAY-1999.  
XX  
PF 09-OCT-1998; 98WO-IB01665.  
XX  
PR 01-SEP-1998; 98GB-0019016.  
PR 06-NOV-1997; 97GB-0023516.  
PR 14-NOV-1997; 97GB-0024190.  
PR 18-NOV-1997; 97GB-0024386.  
PR 27-NOV-1997; 97GB-0025158.  
PR 10-DEC-1997; 97GB-0026147.  
PR 14-JAN-1998; 98GB-0000759.  
XX  
PA {CHIR-} CHIRON SPA.  
XX  
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;  
XX  
DR WPI; 1999-327407/27.  
DR P-PSDB; AAY38561.

Proteins from Neisseria meningitidis and N. gonorrhoeae useful for diagnosis, treatment and prevention of infection

Claim 9; Page 122-123; 524pp; English.

Nucleotide sequences AAZ11972-212358 represent open reading frames (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their fragments, their nucleic acids and antibodies are used for diagnosis, prevention (as vaccines) or treatment of Neisseria infections,

CC such as meningitis, septicaemia and gonorrhea. Both organisms  
CC are closely related. Fragments of the nucleic acids are useful  
CC as hybridisation probes and antisense reagents.  
XX  
SQ Sequence 1344 BP; 329 A; 366 C; 365 G; 284 T; 0 other;

Query Match 91.5%; Score 1229.4; DB 20; Length 1344;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1 atgattaaatacaaaaaggtctaaacctgcctacatcgccgagcagcagcaagtcat 60  
Db 1 atgattaaatacaaaaaggtctaaacctgcctacatcgccgagcagcagcaagtcat 60  
QY 61 tatgacggtccggtcattaccgaagtcggttgcgttggaagaataatccggtatgac 120  
Db 61 tatgacggtccggtcattaccgaagtcggttgcgttggaagaataatccggtatgac 120  
QY 121 cctnagatgaagtcgaaggaagcgatgcgtcaaaaagccaaagtcgctgtttgaagac 180  
Db 121 cctcgtatgaagtcgaaggaagcgatgcgtcaaaaagccaaagtcgctgtttgaagac 180  
QY 181 aaaaagnatccggtggtgtttaccggttcggttcggttcggttcggttcggttcg 240  
Db 181 aaaaagnatccggtggtgtttaccggttcggttcggttcggttcggttcggttcg 240  
QY 241 cgcgcgaaaagcgctacttcagtcggttcggttcggttcggttcggttcggttcg 300  
Db 241 cggtcgaaaagcgctacttcagtcggttcggttcggttcggttcggttcggttcg 300  
QY 301 gagtccgaagcgtacggtccggaagcggttgcgttcggttcggttcggttcggttcg 360  
Db 301 gagtccgaagcgtacggtccggaagcggttgcgttcggttcggttcggttcggttcg 360  
QY 361 aatctgacaaatccggttgcgttcggttcggttcggttcggttcggttcggttcg 420  
Db 361 aatctgacaaatccggttgcgttcggttcggttcggttcggttcggttcggttcg 420  
QY 421 ggcgtgcgtgcgagccgttcggttcggttcggttcggttcggttcggttcggttcg 480  
Db 421 ggcgtgcgtgcgagccgttcggttcggttcggttcggttcggttcggttcggttcg 480  
QY 481 gcagaccctgtgtgtgatcaaaagccggttcggttcggttcggttcggttcggttcg 540  
Db 481 gcagaccctgtgtgtgatcaaaagccggttcggttcggttcggttcggttcggttcg 540  
QY 541 ttgagccgtttgacgagcgtaaaatccatgtgtgtaagcgagctggcgagcggtcg 600  
Db 541 ttgagccgtttgacgagcgtaaaatccatgtgtgtaagcgagctggcgagcggtcg 600  
QY 601 tctgaaaatgctgcaacatcgaaacacatgaattcggttcggttcggttcggttcg 660  
Db 601 tctgaaaatgctgcaacatcgaaacacatgaattcggttcggttcggttcggttcg 660  
QY 661 agtgcagcgacattcatttcattgagccgttcggttcggttcggttcggttcggttcg 720  
Db 661 agtgcagcgacattcatttcattgagccgttcggttcggttcggttcggttcggttcg 720  
QY 721 aattatgaagtgtaattggttcggttcggttcggttcggttcggttcggttcggttcg 780  
Db 721 aattatgaagtgtaattggttcggttcggttcggttcggttcggttcggttcggttcg 780  
QY 781 cgcgtgattgttcggttcggttcggttcggttcggttcggttcggttcggttcggttcg 840  
Db 781 cgcgtgattgttcggttcggttcggttcggttcggttcggttcggttcggttcggttcg 840  
QY 841 ggtgcgaagtcgcaaaatctggttcggttcggttcggttcggttcggttcggttcggttcg 900  
Db 841 ggtgcgaagtcgcaaaatctggttcggttcggttcggttcggttcggttcggttcggttcg 900  
QY 901 tccggttcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttcg 960  
Db 901 tccggttcggttcggttcggttcggttcggttcggttcggttcggttcggttcggttcg 960







Db 301 gaattcgaaagcgtacgtacgtgaagcgtggaataattgagcagcgaataaagtgcgcgc 360  
 QY 361 aatctgatccaaatccggtttgtgactgcgtgcgttancgctccgttcagcaaaatccct 420  
 Db 361 aacctgattcaatcaggtattgactgcgttcgaccgcgtccgttcagcaaaatccct 420  
 QY 421 gcgctgatgcgagcgttcgctccatcttcgtcaatgcgatgacacaaatccgctnccg 480  
 Db 421 gccgtatgcgagcgttcgctccatcttcgtcaatgcgatgacacaaatccgctnccg 480  
 QY 481 gcagaccctgtggttgatcaaaagcgcngcngatcttcagacgngtntgtgtga 540  
 Db 481 gcgaccctcaggttcacatacgaagcgcgcgagcgttcacacgttcgaagcgtgtgtga 540  
 QY 541 ttgagcgtttgacgagcgttaaatccatgtgtgtaagcagcgtgcgcagacgtgcgc 600  
 Db 541 ttgagcgttcgacgagcgttaaatccatgtgtgtaagcagcgtgcgcagacgtgcgc 600  
 QY 601 tctgaaatgctgcacacgtgaacacacatgaattcggcgcgcgtccgctggtttg 660  
 Db 601 tctgaaatgctgcacacgtgaacacacatgaattcggcgcgcgtccgctggtttg 660  
 QY 661 agtgacgcgcacatcattcattgagcgcgtgcgtgcgaacacaaacgtttgacacac 720  
 Db 661 agtgacgcgcacatcattcattgagcgcgtgcgtgcgaacacaaacgtttgacacac 720  
 QY 721 aattatcaagatgaattgccatcgcgcgtttgttgcaacagcgcgtctgaacacag 780  
 Db 721 aattatcaagatgaattgccatcgcgcgtttgttgcaacagcgcgtctgaacacag 780  
 QY 781 cgcgtgattgtttggtgttctcaagtcacaaacacgcgcgtcttcgttaccgtttg 840  
 Db 781 cgcgtgattgtttggtgttctcaagtcacaaacacgcgcgtcttcgttaccgtttg 840  
 QY 841 ggtgcgaagatgcgaattactgcgcgcgaattgttaacgagcgcgttcgaacacacgcgttatt 900  
 Db 841 ggtgcgaagatgcgaattactgcgcgcgaattgttaacgagcgcgttcgaacacacgcgttatt 900  
 QY 901 tccggttcggtattgaagcgcgcgtattacacagcgcgcgtattttgggacgtac 960  
 Db 901 tccggttcggtattgaagcgcgcgtattacacagcgcgcgtattttgggacgtac 960  
 QY 961 cacaatcagattcgttattgcgaagcgcgcgaagcgcgcgttcggtggttgcg 1020  
 Db 961 cacaatcagattcgttattgcgaagcgcgcgaagcgcgcgttcggtggttgcg 1020  
 QY 1021 ccgagcgcgcgaataactcctacgcgtacgacctcgcgcgttcctgaaacacaa 1080  
 Db 1021 ccgagcgcgcgaataactcctacgcgtacgacctcgcgcgttcctgaaacacaa 1080  
 QY 1081 ctcttcgaagttcagacagcgcgtcaacggtggcgcgcgcgttcggtggttact 1140  
 Db 1081 ctcttcgaagttcagacagcgcgtcaacggtggcgcgcgcgttcggtggttact 1140  
 QY 1141 tacgagcgcgttaattgcgcgttagacatcctcctaccctggtttgagcgtatttaactgc 1200  
 Db 1141 tacgagcgcgttaattgcgcgttagacatcctcctaccctggtttgagcgtatttaactgc 1200  
 QY 1201 ggcgataccagcgcgcgaagcatttgggttggattggaattgacgaagacacgtcgcgt 1260  
 Db 1201 ggcgataccagcgcgcgaagcatttgggttggattggaattgacgaagacacgtcgcgt 1260  
 QY 1261 ttgtcagcttcgttcgcccgcgaataacgaatgngccgctgtgttgtaaggtgctg 1320  
 Db 1261 ttgtcagcttcgttcgcccgcgaataacgaatgngccgctgtgttgtaaggtgctg 1320  
 QY 1321 gaaacnttgagaagaaggtcga 1344  
 Db 1321 gaaacnttgagaagaaggtcga 1344

RESULT 6  
 AAT42063

ID AAT42063 standard; DNA; 1830121 BP.  
 XX AAT42063;  
 AC 14-SEP-1999 (first entry)  
 DT Haemophilus influenzae complete genome sequence.  
 DE Genome; bacterium; Haemophilus influenzae; computer readable medium;  
 KW expression modulating fragment; regulation; gene expression; vector;  
 KW organism; open reading frame; ORF; ds.  
 XX Haemophilus influenzae.  
 OS Haemophilus influenzae.  
 XX W09633276-A1.  
 XX 24-OCT-1996.  
 XX 22-APR-1996; 96WO-US05320.  
 XX 07-JUN-1995; 95US-0487429.  
 PR 21-APR-1995; 95US-0426787.  
 PR 07-JUN-1995; 95US-0476102.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;  
 XX WPI; 1996-485782/48.  
 DR Haemophilus influenzae Rd genome recorded on computer readable  
 PT medium - useful for identifying commercially important nucleic acid  
 PT fragments by homology searching  
 XX Claim 1; Page 77.2-77.1091; 1291pp; English.  
 PS This sequence represents the complete genome sequence of the bacterium  
 CC Haemophilus influenzae strain Rd. The invention relates to a computer  
 CC readable medium (CRM) having recorded upon it the complete H. influenzae  
 CC nucleotide sequence (1), a representative fragment of (1) or a nucleotide  
 CC sequence at least 99% identical to (1). By providing the full-length  
 CC genomic sequence in a computer readable form, it is possible to identify  
 CC commercially important nucleic acid fragments and expression modulating  
 CC fragments (EMEs) of the Haemophilus genome. The EMFs can be used to  
 CC regulate the expression of a nucleic acid molecule. Vectors and altered  
 CC organisms comprising the predicted ORFs can be used to produce any of the  
 CC polypeptide fragments of the H. influenzae Rd genome.  
 XX Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Query Match 39.2%; Score 527.2; DB 17; Length 1830121;  
 Best Local Similarity 62.5%; Pred. No. 6e-150;  
 Matches 842; Conservative 0; Mismatches 500; Indels 5; Gaps 2;  
 QY 1 atgattataatacaaaaagggtctaaacctgccatcgccgagcagccgagcaagtcatt 60  
 Db 179007 atgattataatacaaaaagggttggatctcccaattcggaacacagcaagtaac 179066  
 QY 61 tatgacggcccgctcattaccgaagtcgctgttcgcaagaataatgcggtatgcgc 120  
 Db 179067 catagcggcaacgcgtgtgaatcaagtgcgattctagtgaggagatgtggtgacgc 179126  
 QY 121 cctctngatgaaagtcaagaaggcgatccgctcaaaaagcccaagtgctgtttgaagac 180  
 Db 179127 ccttcaatgaagggtgcgcgaaggcgatgttgaaaaaaggctcaagtacttttgaagac 179186  
 QY 181 aaaaagnatccggcgctgtgtttaccgcgcngtttcagcgaataatgcgcgcacatccat 240  
 Db 179187 aaaaaaatcctggtgttaatttttacagccctcgaagcgttaccactgcaatcaat 179246  
 QY 241 cgcgcgcaaaaagcgcgctactcagtcggtgctgattgcgttggaaggcaacgacgaatc 300



Db 179247	cggtgcgaacacgctgattacaactctgtgtttaataatagaaaggtgatgaaaaatc	179306	Db 180325	ttagataagattgagaaggaaggttaa	180351
QY 301	gagttcgaacgtacgcgcgcgaagcgttggcaaaacttgaagcgcgangaantnnngnc	360	RESULT 7		
Db 179307	acttttgcaaatatagacagaacaattgaatacactttcttctgaacaagtataaaca	179366	AAZ12025	AAZ12025 standard; DNA; 474 BP.	
QY 361	aatctgatccaatccoggtttgttgactgcgtgcgtanccgtccgtttcagcaaaaatccct	420	XX	AAZ12025;	
Db 179367	aaacttgagaactctggttatggacgcgcatctacgtactcgtccatttagtaaaagtcca	179426	XX	08-OCT-1999 (first entry)	
QY 421	gcgctgatgcgagcgcgttcgcatcttcgtaactgcgatgacaccaaaccgctgngc	480	XX	Neisseria meningitidis partial ORF22 sequence.	
Db 179427	tcattgaaaggaagcactctctattttgttaattgcgatgataccaatcattagca	179486	DE	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
QY 481	gcagaccctgtggttatgatcaaaagcgcgncgngatttccagcngangntgtggtga	540	KW	treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.	
Db 179487	gcgataccctctgctgtattaaagaataattcgcaagacttcaactaagcgttttaacogta	179546	XX	Neisseria meningitidis.	
QY 541	ttgagcgcgttgaccagagcgttaa---aatccatgtgtgtgaagcagctggcgcaacgtg	597	OS	Neisseria meningitidis.	
Db 179547	ttaagtcgtctattcccttcaaacacattacacttattgaagcagcggattctctaatt	179606	XX	WO9924578-A2.	
QY 598	ccgtctgaaaaatgctccaaacatcgcaaacacacatgaattcggcgcgcgcgcgcgcgt	657	XX	20-MAY-1999.	
Db 179607	ccaaactgcgtatcttgaaaaatttacagattcatgttttacttgcgtttcatccagctgtg	179666	XX	09-OCT-1998; 98WO-IB01665.	
QY 658	ttgattggcagcacaattcatttcaattgagcgcgtgcgtgcgaacaaacacccgtttggacc	717	XX	01-SEP-1998; 98GB-0019016.	
Db 179667	cttctagggacgcatactcacttattgtatcctgcgtgtattcaaaaactgtatggcat	179726	PR	06-NOV-1997; 97GB-0023516.	
QY 718	ataaattacaagatgaattgcacgcgcgcgtttgttgcaacagcgcgcgtctggaaccc	777	PR	14-NOV-1997; 97GB-0024190.	
Db 179727	ataaattacaagatgtagctgctgtaggttaaatattttacaacagcgaactttattca	179786	PR	18-NOV-1997; 97GB-0024386.	
QY 778	gagcgcgtgattgctttgggtgttctcaagtcacacaaacccacccctcttgcgtaccgtt	837	PR	27-NOV-1997; 97GB-0025158.	
Db 179787	gaacagtaattctctctgagctcctcaagtgaagaaacccctgttttagtactacaca	179846	PR	10-DEC-1997; 97GB-0026147.	
QY 838	ttgggtgcgaagtatgcgaataactactgcggcgaattgttgacacgaacccgcgtg	897	PR	14-JAN-1998; 98GB-0000759.	
Db 179847	atcgtgtcaattctctcatttaaccccaaaataagttgaagtcaggtgtaaaatcgtgta	179906	XX	(CHIR-) CHIRON SPA.	
QY 898	attccggttcggtattgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	957	XX	Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;	
Db 179907	attctgcgtcagttctgttggtcaaatagcaaaagattcacatgattatttagtgcgc	179966	XX	WPI; 1999-327407/27.	
QY 958	taccacaatcagatttccgttatogaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	1017	DR	P-PSDB; AAY38560.	
Db 179967	tatctttgcaagtacagtaactcgtgaaggtgaatgaagaaatgttttgggttgatt	180026	PT	Proteins from Neisseria meningitidis and N. gonorrhoeae useful for	
QY 1018	gc	1077	PT	diagnosis, treatment and prevention of infection	
Db 180027	atgcacaagcccaataatctctgactcgtacgcgtacgcgcgcgcgcgcgcgcgcgcgcgc	1137	XX	Claim 9; Page 122; 524pp; English.	
QY 1078	aaactctcaagttcagcagc	1197	XX	Nucleotide sequences AAZ11972-212358 represent open reading frames	
Db 180085	aaattgttaattttactacttcagaaatggcgcgaacgtgcaatggcgcaatcgcgt	180144	CC	(ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode	
QY 1138	acttacagc	1197	CC	antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their	
Db 180145	agctatgagc	180204	CC	fragments, their nucleic acids and antibodies are used for diagnosis,	
QY 1198	gtcgc	1257	CC	prevention (as vaccines) or treatment of Neisseria infections,	
Db 180205	gtggc	180264	CC	such as meningitis, septicaemia and gonorrhea. Both organisms	
QY 1258	gctttgtcagcttctgctgc	1317	CC	are closely related. Fragments of the nucleic acids are useful	
Db 180265	gcggtatgttcttctgcttgc	180324	CC	as hybridisation probes and antisense reagents.	
QY 1318	ctggaacacncttgagaaggaagcgtga	1344	XX	Sequence 474 BP; 122 A; 130 C; 130 G; 91 T; 1 other;	







Db	84964	CGTGGCGACGACGAGATATTGAAAGCTTGGTCATCGCGCTGATCCAAATGCAAGTGAA	84905
QY	295	gaaatcgaagtctgaacgctacgcgcgcgaagcgtttggcaaaacttaagcgcgcgangaantn	354
Db	84904	GAAATCACTTTTAATGCTTATCGCTCTCGAAGATTTGTCATAGCTGATCGCAAGACAGTT	84845
QY	355	ngnngcaatctgatccaatccgggttttgagctgcgcgtgtancgcgtccgttcagcaaa	414
Db	84844	GTTGACGACTTGGTCGCTTCAGGTGAATGGACCGCCTACGCACTCGTCTTTTAGCCGT	84785
QY	415	atccctgcgcgtgatgcgcgacgcgttcgcgcattcttcgcaatgcgatggacacaatccg	474
Db	84784	ACACGAGAAATTAATAGCATACCTTCTCGCATTTTGTGACGGCAACAGATACCAATCCT	84725
QY	475	ctnngcgcgacccctgtgttgatcaaaagaagcngcngangatttcagacganganctg	534
Db	84724	TTGGCTGTGACCCCTGCCCGATTATTATGAGAGAAATACAGCATTTAATGATGCGCTTG	84665
QY	535	ctggtattgagccgtttgaccgagcgtataatccattgtgtaagcagctggcgacac	594
Db	84664	GCAATTTTATCAACGCTATCC--CCAAAACCATGTCTGTCTGTCATGGTGATATTGACCT	84608
QY	595	gtgcgcgtgaaaaatgctgccaac-----atcgaaacacatgaattcggcgcccg	645
Db	84607	GAGACCTTAAGTAAACCGCTGACGGTAAACGATACCATCTATGAAGGGTTTACTGGTGA	84548
QY	646	caticgcgcggttgagtggcacgacatcatttcattgatgcgcggtcggtgcacaacaa	705
Db	84547	CATCCAGCAGGCAACGCAGGTACGCACATTCACCTTTTACACCTTTGGCAGCAGGAGC	84488
QY	706	accgtttgaccatcaattatacagatgtaatggcatcggcgctgtttgttgaacagcc	765
Db	84487	TTAGTTTGGACGATTAAATATCAAGATGTCATCGCCATTTGTAAGCTATTTACCACAGT	84428
QY	766	cgtctgaacacgcgcgcgtgatgtgtttgggttggttctcaagtcacaaacccgcctc	825
Db	84427	CGCATTTACACAGACGCTCTCATAGTTTGGCAGGTCCACAGTGGCCAAATCCTGCCTG	84368
QY	826	ttgcgtaccgctttgggtgcgaagtatcgcaattactgcgggcgaattggtgaagca	885
Db	84367	GTGCGTACCACCTCGAGGTGCTGATTTGACCGAATTTGACCCAGGTGAATCAACGGCCAC	84308
QY	886	gcaaccgcgctgattccggttcggtattgaacgcgcgattacacaaagcgcgcagcat	945
Db	84307	GATAACCGGTGTCATTTTCAGGCTCTGTTCTATCAGTTCGCACAGCTGTTGTCACGCA	84248
QY	946	tatttggcagctaccacaatcagatttcgcgttatcgaagaagcgccgcagcaagagctg	1005
Db	84247	TTTTTTGGCGCGTTCATTAATCAAGTCTCTGTACTTACCAGAGGTGCTGACAGCCCTGCA	84188
QY	1006	tlcgcgctgggttcgcgcgcgcgcgaacatactccatcacgcgtacgacctcggccat	1065
Db	84187	CTTCACCTTTTTCACACGAGGGCAATTCGTTTTCAAAATTCGCAATTTATATTTACAG	84128
QY	1066	ttcctgaaacaaactcttcaagttccacgacgcgcgtcaacgcgttgccgacgcgcctatg	1125
Db	84127	TTTTTTAAGGCAAAAATACGATTTTACGACCAGTTACCAACGGCTCGCTCGTCCCATG	84068
QY	1126	gtgcgcgattgtaacttcagcgcgcgtaatgccgcgtagacatcctgcctacctgctttg	1185
Db	84067	GTACCAATCGGCTTTTGTAAACTGTCATGCCCTCAAGACTATCTACCTACCCAGCTACTG	84008
QY	1186	cgcgatttaatcgtcgcgcgataccgacagcgcgcgaagcatttggttctgttgaaattgac	1245
Db	84007	CGTGCTTTTGATTTGTCGAAGATATCATCAGTGCGGTTGCAATTTGGGTGCATTTGAGCTGAT	83948
QY	1246	gaagaagacctcgcgtttgtgcagcttcgctcgtcccgggcgaatacgaatatangccgcgtg	1305
Db	83947	GAGAAGATTGGCACTGTGTACTTTTCGTATCGCCAGGCAATATGAGTTCGGTGATATT	83888
QY	1306	ttgcgttaaggtgctggaaacconttgaaagaaggctga	1344

[illegible]

































CC Mycobacterium tuberculosis or related Mycobacterium by determining the  
CC nucleotide sequence of the first strain at positions in the complete  
CC sequence of the genome that correspond to positions that differ in the  
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (AA199883) and  
CC H37Rv (AA199882). The method is useful for evaluating strain variation of  
CC M. tuberculosis and has valuable application in the fields of  
CC tuberculosis genetics, epidemiology, patient treatment and epidemic  
CC monitoring.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at [seqdata.uspto.gov/sequence.html?DocID=6294328B1](http://seqdata.uspto.gov/sequence.html?DocID=6294328B1).  
XX  
SQ Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Query Match 2.7%; Score 35.8; DB 22; Length 4411529;  
Best Local Similarity 55.1%; Pred. No. 98;  
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 1044 cagcgtacgaccctcgccattctctgaaacaaactctcaagttcacgacgacct 1103  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 4110332 CACCCTAACACATCGGCCATGACAAAGACGGTGACCTCATCGGTTCACTTCGCCCA 4110273  
QY 1104 caacggtggcagcccgccatgtgcccgtattgtacttacgagcgcgtaatgcgctaga 1163  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 4110272 CTGTGGGGCCCGCGCGCAGATCGCCGCTGAGGCTGCCGAGAACGGCGGCCGTAGG 4110213  
QY 1164 cactctg 1170  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 4110212 CATCATG 4110206

RESULT 33  
AA08045  
ID AA08045 standard; cDNA; 2558 BP.  
XX  
AC AA08045;  
XX  
DT 19-JUN-2000 (first entry)  
XX  
DE f-Therien rubella virus cDNA clone sequence SEQ ID NO:3.  
XX  
KW Rubella virus; infectious; infectious; chimeric construct; anti-viral;  
KW vaccine; birth defect; autoimmune disease; ss.  
XX  
OS Rubella virus.  
XX  
PN US6054573-A.  
XX  
PD 25-APR-2000.  
XX  
PF 02-SEP-1997; 97US-0999733.  
XX  
PR 28-JUN-1991; 91US-0722334.  
PR 19-JUL-1993; 93US-0093453.  
PR 02-JUN-1995; 95US-0459041.  
XX  
PA (UVGE-) UNIV GEORGIA STATE.  
XX  
PI Abnathay ES, Pougatchev K, Frey TK;  
XX  
DR WPI; 2000-328366/28.  
XX  
PT Highly infectious rubella virus clones useful for developing a rubella  
PT vaccine that can be safely administered to pregnant and older women  
PT without risk of birth defects and autoimmune disease  
XX  
PS Claim 2; Column 21-24; 17pp; English.  
XX  
CC The present invention describes nucleic acid molecules (AA08044 and  
CC AA08045) which are fragments of the f-Therien rubella virus genome.  
CC AA08044 and AA08045 are used to replace the corresponding fragments of  
CC infectious rubella virus cDNA clone with low specific infectivity

CC referred to as Robol02 to create a chimeric construct with high  
CC specific infectivity, Robo302. The highly infectious rubella virus  
CC clones are useful as molecular biology tools for studying rubella virus  
CC and can be used for developing recombinant vaccines against rubella.  
CC The rubella vaccines developed can be safely administered to pregnant  
CC and older women without risk of birth defects or autoimmune disease.  
XX  
SQ Sequence 2558 BP; 379 A; 991 C; 805 G; 383 T; 0 other;

Query Match 2.6%; Score 35.6; DB 21; Length 2558;  
Best Local Similarity 45.4%; Pred. No. 3.2;  
Matches 122; Conservative 0; Mismatches 147; Indels 0; Gaps 0;  
QY 55 gtcattatgacggcgccctcattaccgaagtgcggtgttggcgaagaatatgccggt 114  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1142 ggcactgtgcgcgacgacgacgaggggtgcccaggcgctactacgacgacctcgagtg 1201  
QY 115 atgcgccccctngatgaaagtcaaggaagcgatgcgtcaaaaagcccaagtgctgttt 174  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1202 cgcgcctcggggatgacgcatggcccgggcgccctcgcatcagtcacacgcccctgc 1261  
QY 175 gaagacaaaagatnaccgggggtgtgtttaccgcccngtttcagcaaaatccgcgc 234  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1262 aaaggcccttacaatcaggggtatggaacatggcgcgagcgctggcaagactaccgc 1321  
QY 235 atccatcgcggaagcgcgctacttcagtcgctgattgcgttgaaaggcaagac 294  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1322 atccctgctgcttcacgcgcgagacatttactgtctcccccaaatgcgctctgcac 1381  
QY 295 gaaatcgagttcgaaacgctacgcgcga 323  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1382 gagatccaggccaaactccgcgcgcga 1410

RESULT 34  
AAQ97686  
ID AAQ97686 standard; RNA; 9757 BP.  
XX  
AC AAQ97686;  
XX  
DT 27-FEB-1996 (first entry)  
XX  
DE Infectious rubella virus RNA.  
XX  
KW Rubella; vaccine; mutant; epitope; virus; autoimmune disease;  
KW pregnancy; foetal infection; vector; plasmid; ss.  
XX  
OS Rubella virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 41..6658  
FT /\*tag= a  
FT /product= N-terminal transcript.  
FT misc\_difference 2261..2263  
FT /\*tag= b  
FT /transl\_except= CGU encodes Ala.  
FT misc\_difference 6605..6607  
FT /\*tag= c  
FT /transl\_except= GUC encodes asparagine or aspartic acid.  
FT misc\_difference 8460..8462  
FT /\*tag= d  
FT /transl\_except= CUG encodes Proline.  
FT misc\_difference 8463..8465  
FT /\*tag= e  
FT /transl\_except= CUC encodes Cysteine.  
FT misc\_difference 9075..9077  
FT /\*tag= f  
FT /transl\_except= UGG encodes Methionine.  
XX  
XX US5439814-A.  
XX  
XX 08-AUG-1995.





QY 1291 gaatanggcc 1300  
DB 826 ggcacgcacc 835

RESULT 38  
AAS57519  
ID AAS57519 standard; cDNA; 296 BP.  
AC AAS57519;  
XX  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX cDNA #195 encoding portion of a human colon tumour protein.  
XX  
XX Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200173027-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 22-MAR-2001; 2001WO-US09246.  
XX  
XX 24-MAR-2000; 2000US-191597P.  
XX 04-MAY-2000; 2000US-202024P.  
XX 05-MAY-2000; 2000US-202189P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Meagher MJ, Xu J, King GE;  
XX  
XX WPI: 2001-611627/70.  
XX  
XX New colon tumour proteins and related nucleic acid, useful for  
XX treatment, prevention, diagnosis and monitoring of cancer.  
XX  
XX Claim 4; Page 90; 299pp; English.  
XX  
XX Th present invention relates to the isolation of novel cDNA sequences  
XX encoding for at least an immunogenic portion of human colon tumour  
XX proteins. The sequences of the invention are useful in pharmaceutical  
XX compositions and vaccines for the prevention and treatment of cancers  
XX such as colon cancer. They are also useful for the diagnosis and  
XX monitoring of such cancers. Antibodies to the colon tumour proteins  
XX and antigen presenting cells that express polynucleotides encoding  
XX colon tumour proteins can be used to inhibit the development of  
XX cancers. T-cells that react specifically with colon tumour proteins  
XX are useful for removing tumour cells from samples (e.g. blood) and  
XX for cancer treatment. The polynucleotides sequences are also useful in  
XX gene therapy. AAS57325-AAS5880 represent the cDNA sequences of the  
XX invention that encode for portions of human colon tumour proteins.  
XX  
XX Sequence 296 BP; 66 A; 86 C; 88 G; 51 T; 5 other;

Query Match 2.6%; Score 35; DB 23; Length 296;  
Best Local Similarity 57.8%; Pred. No. 1.7;  
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1239 attggcgaagacacctgttggcagcttcgtcccgccggaacacgaatgagg 1298  
DB 74 angggccgcgagagtgaggccctgcgctgcctgcacacccgggcacattcgaggatg 133

QY 1299 ccgcgtgttcgtaaggtctggaacnnttgagaaggaagg 1340  
DB 134 ccaccggaagtgcgaagagctgttcccatcagatggagg 175

RESULT 39  
AAV23811

ID AAV23811 standard; DNA; 1935 BP.  
XX  
XX AAV23811;  
XX  
XX 29-JUL-1998 (first entry)  
XX  
XX Nitrosomonas dnaK gene.  
XX  
XX Heatshock promoter; dnaK gene; stress-sensing gene; oxidative stress; ss.  
XX  
XX Nitrosomonas europaea.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1935  
XX /\*tag= a  
XX  
XX JP10108678-A.  
XX  
XX 28-APR-1998.  
XX  
XX 07-OCT-1996; 96JP-0266320.  
XX  
XX 07-OCT-1996; 96JP-0266320.  
XX  
XX (KURK ) KURITA WATER IND LTD.  
XX  
XX WPI: 1998-304973/27.  
XX P-PSDB; AAW54020.  
XX  
XX New heat-shock promoter from Nitrosomonas species - useful for, e.g.  
XX measuring oxidative stress caused by ammonia  
XX  
XX Example 1; Page 10-13; 18pp; Japanese.  
XX  
XX This sequence represents a the dnaK gene of Nitrosomonas  
XX europaea. The heatshock promoter (HP) of the amplified DNA is an example  
XX of the promoter of the invention. The HP can be used in a stress-sensing  
XX gene comprising: (a) a fused DNA fragment comprising HP; and (b) a DNA  
XX fragment positioned downstream of the fragment of (a), comprising a  
XX structural gene encoding a protein for detecting a gene expression. A  
XX microorganism carrying the stress-sensing gene may be used to measure the  
XX oxidative stress caused by ammonia, by culturing it and measuring the  
XX expression of the stress-sensing gene. The method can measure the  
XX oxidative stress easily and rapidly.  
XX  
XX Sequence 1935 BP; 574 A; 442 C; 546 G; 373 T; 0 other;

Query Match 2.6%; Score 35; DB 19; Length 1935;  
Best Local Similarity 50.3%; Pred. No. 4.3;  
Matches 77; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 221 gcaaatcgccgcacatccatcgccgcaaaagcgctacttcagtcggtcgatgacg 280  
DB 563 gcaaatcgccgcacatccatcgccgcaaaagcgctacttcagtcggtcgatgacg 622

QY 281 ttgaaggcaacgacgaaatcgagttcgacgctacgcccgcgagcgttggaactaa 340  
DB 623 cagaagtcgaagggcgaacacagttcgagtgctggaacacacacacacacacacac 682

QY 341 gcggcggaagantngnngcaatctgatccaat 373  
DB 683 gcggcgagattcgactccacgagatgacgaat 715

RESULT 40  
AAS86095  
ID AAS86095 standard; cDNA; 323 BP.  
XX  
XX AAS86095;  
XX  
XX 13-FEB-2002 (first entry)  
XX







Db	1068	gctgctgtccaagccaatgccgaggtggtgcccagtgagagagcaagtagcagacagatgc	1127
Qy	150	cgtaaaaaagggcaagtgctgtttgaagcaaaaaaatccggcggtgtgtttaccgc	209
Db	1128	catccagagaccgagagctggagagggccaaaaaaagctggcactggcggtgcagga	1187
Qy	210	gcengtttcaggcaaaatcgccgcacac	237
Db	1188	ggcagagagggcggtggaggctgccaac	1215
RESULT 44			
ID	AAS73416/c		
XX	AAS73416 standard; cDNA; 1050 BP.		
AC	AAS73416;		
DT	13-FEB-2002 (first entry)		
DE	DNA encoding novel human diagnostic protein #9220.		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
OS	Homo sapiens.		
XX	WO200175067-A2.		
PN	11-OCT-2001.		
PD	30-MAR-2001; 2001WO-US08631.		
PF	31-MAR-2000; 2000US-0540217.		
PR	23-AUG-2000; 2000US-0649167.		
XX	(HYSP-) HYSEQ INC.		
PA	Drmanac RT, Liu C, Tang YT;		
PI	WPI; 2001-639362/73.		
XX	P-PSDB; ABG09229.		
XX	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity -		
XX	Claim 1; SEQ ID No 9220; 103pp; English.		
PS	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
CC	and gene mapping, and in recombinant production of (II). The		
CC	polynucleotides are also used in diagnostics as expressed sequence tags		
CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
CC	to restore normal activity of (II) or to treat disease states involving		
CC	(II). (II) is useful for generating antibodies against it, detecting or		
CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
CC	a food supplement. (II) and its binding partners are useful in medical		
CC	imaging of sites expressing (II). (I) and (II) are useful for treating		
CC	disorders involving aberrant protein expression or biological activity.		
CC	The polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		
CC	and to produce other types of data and products dependent on DNA and		
CC	amino acid sequences. AAS64197-AAS94564 represent novel human		
CC	diagnostic coding sequences of the invention.		
CC	Note: The sequence data for this patent did not appear in the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 1050 BP; 252 A; 299 C; 264 G; 235 T; 0 other;		
XX	50		
Query Match			
2.6%; Score 34.4; DB 23; Length 1050;			
Best Local Similarity 45.1%; Pred. No. 4.8;			
Matches 110; Conservative 0; Mismatches 134; Indels 0; Gaps 0;			
Qy	128	tgaagtaaggaaggcgatgccgtcaaaaaagggccaaagtgtgttgaagacaaaaaagn	187
Db	608	TGAGTGACAAAAATCAGGCTGAGCGCGGATCGCGCAACTGACGATCAGGTCAACACAGT	549
Qy	188	atccggggcggtgtttaccgcgccngtttcaggcaaaatcgccgcacatccatcgcgcg	247
Db	548	AACTCCACTTCCTGGGTGTCGAGACTGGCGGTGGTTCATCGAGGATCAGCATTGGCA	489
Qy	248	aaaagcgcgtacttcagtcggtcgatgccgttgaaggcaaacgacgcaaatcgagttcg	307
Db	488	GAGAGATCGATAGCCGACAAATAGCGACGATTTGCTGATCGGACTGAAAAGCGTTG	429
Qy	308	aacgctacgccccgaagcgttgcaaaacttaagcgcgcgaangannnnngcaactcga	367
Db	428	AGCGGTTTCGCACGTCGAGGGAGAAACCGTAAGATGCCATCAGTTCGGTGGCGGCTTT	369
Qy	368	tcca 371	
Db	368	tcca 365	
RESULT 45			
ID	AAS80144/c		
XX	AAS80144 standard; cDNA; 1050 BP.		
AC	AAS80144;		
DT	13-FEB-2002 (first entry)		
DE	DNA encoding novel human diagnostic protein #15948.		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
OS	Homo sapiens.		
XX	WO200175067-A2.		
PN	11-OCT-2001.		
PD	30-MAR-2001; 2001WO-US08631.		
PF	31-MAR-2000; 2000US-0540217.		
PR	23-AUG-2000; 2000US-0649167.		
XX	(HYSP-) HYSEQ INC.		
PA	Drmanac RT, Liu C, Tang YT;		
PI	WPI; 2001-639362/73.		
XX	P-PSDB; ABG15957.		
XX	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity -		
XX	Claim 1; SEQ ID No 15948; 103pp; English.		
PS	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
CC	and gene mapping, and in recombinant production of (II). The		
CC	polynucleotides are also used in diagnostics as expressed sequence tags		
CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
CC	to restore normal activity of (II) or to treat disease states involving		
CC	(II). (II) is useful for generating antibodies against it, detecting or		
CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
CC	a food supplement. (II) and its binding partners are useful in medical		
CC	imaging of sites expressing (II). (I) and (II) are useful for treating		
CC	disorders involving aberrant protein expression or biological activity.		
CC	The polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		
CC	and to produce other types of data and products dependent on DNA and		
CC	amino acid sequences. AAS64197-AAS94564 represent novel human		
CC	diagnostic coding sequences of the invention.		
CC	Note: The sequence data for this patent did not appear in the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 1050 BP; 252 A; 299 C; 264 G; 235 T; 0 other;		
XX	50		



